THE RANK OF $n \times n$ MATRIX MULTIPLICATION IS AT LEAST $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$

ALEX MASSARENTI AND EMANUELE RAVIOLO

Abstract. We prove that the rank of the $n \times n$ matrix multiplication is at least $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$. The previous bounds were $3n^2 - 4n^{\frac{3}{2}} - n$ due to Landsberg [L] and $\frac{5}{2}n^2 - 3n$ due to Bläser [B]. Our bound improves the previous bounds for any $n \geqslant 24$.

Contents

Introduction

The multiplication of two matrices is one of the most important operations in pure mathematics and applied sciences. To determine the complexity of matrix multiplication is a major open question in algebraic complexity theory. A measure of the complexity of matrix multiplication, and of tensors in general, is the *rank*.
Roughly speaking, the rank is the minimum number of scalar multiplications among the entries of two $n \times n$ matrices to obtain the product. The standard algorithm requires n^3 scalar operations, but *V. Strassen* showed that that such algorithm is not optimal [S]. In this paper we are concerned with lower bounds on the rank of matrix multiplication. The first lower bound $\frac{3}{2}n^2$ was proved by *V. Strassen* [S1] and then improved by *M. Bläser* [B], who found the lower bound $\frac{5}{2}n^2 - 3n$.
Recently *J.M. Landsberg* [L], building on work with *G. Ottaviani* [LO], found the new lower bound $3n^2 - 4n^{\frac{3}{2}} - n$. The core of Landsberg's argument is the proof of the Key Lemma [L, Lemma 4.3]. In this paper we improve the Key Lemma and obtain the following lower bounds for matrix multiplication.

Theorem 0.1. *Let $p \leqslant \frac{n}{2}$ be a natural number. Then*

$$\operatorname{rk}(M_{n,n,m}) \geqslant (1 + \frac{p}{p+1})nm + n^2 - (2p+3)n. \tag{0.1}$$

Date: November 28, 2012.
1991 *Mathematics Subject Classification.* Primary 14Q20; Secondary 13P99, 68W30.
Key words and phrases. Tensors, Matrix Multiplication, Complexity Theory.

For example, when $\sqrt{\frac{n}{2}} \in \mathbb{Z}$, *taking* $p = \sqrt{\frac{n}{2}} - 1$, *we get*

$$\mathrm{rk}(M_{n,n,m}) \geqslant 2nm + n^2 - 2\sqrt{2}nm^{\frac{1}{2}} - n.$$

When $n = m$ *we obtain*

$$\mathrm{rk}(M_{n,n,n}) \geqslant (3 - \frac{1}{p+1})n^2 - (2p+3)n. \tag{0.2}$$

This bound is maximized when $p = \lceil \sqrt{\frac{n}{2}} - 1 \rceil$ *or* $p = \lfloor \sqrt{\frac{n}{2}} - 1 \rfloor$, *hence when* $\sqrt{\frac{n}{2}} \in \mathbb{Z}$ *we have*

$$\mathrm{rk}(M_{n,n,n}) \geqslant 3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - n.$$

In general we have the following bound

$$\mathrm{rk}(M_{n,n,n}) \geqslant 3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n. \tag{0.3}$$

The bound (0.3) improves Bläser's one, $\frac{5}{2}n^2 - 3n$, for $n \geqslant 32$. Nevertheless, when $p = 2$, the bound in (0.2) becomes $\frac{8}{3}n^2 - 7n$, which improves Bläser's one for every $n \geqslant 24$. Compared with Landsberg's bound $3n^2 - 4n^{\frac{3}{2}} - n$, our bound (0.3) is better for $n \geqslant 3$.
Our strategy is the following. We prove Lemma 3.2, which is the improved version of [L, Lemma 4.3], using the classical identities for determinants of Lemma 1.1 and Lemma 1.2, to lower the degree of the equations that give the lower bound for border rank for matrix multiplication. Then we exploit this lower degree as Bläser and Landsberg did.
The paper is organized as follows. In Section 1 we give the basic definitions and explain the geometric meanings of the notions of rank and border rank in terms of secant varieties of Segre varieties. Section 2 is devoted to the Landsberg-Ottaviani equations [LO]; we present them as rephrased in [L]. Finally in Section 3 we improve the Key Lemma [L, Lemma 4.3] and prove Theorem 0.1.

1. Preliminaries and Notation

Let V, W be two complex vector spaces of dimension n and m. The contraction morphism

$$V^* \otimes W \to \mathrm{Hom}(V, W),\ \sum_{i,j} f_i \otimes w_j \mapsto L_{f_i, w_j},$$

where $L_{f_i,w_j}(v) = \sum_{i,j} f_i(v)w_j$, defines an isomorphism between $V^* \otimes W$ and the space of linear maps from V to W.
Then given three vector spaces A, B, C of dimension a, b and c we can identify $A^* \otimes B$ with the space of linear maps $A \to B$, and $A^* \otimes B^* \otimes C$ with the space of bilinear maps $A \times B \to C$. Let $T : A^* \times B^* \to C$ be a bilinear map. Then T induces a linear map $A^* \otimes B^* \to C$ and may also be interpreted as:

- an element of $(A^* \otimes B^*)^* \otimes C = A \otimes B \otimes C$,
- a linear map $A^* \to B \otimes C$.

Rank and border rank of a bilinear map. The *rank* of a bilinear map $T : A^* \times B^* \to C$ is the smallest natural number $r := \mathrm{rk}(T) \in \mathbb{N}$ such that there exist $a_1, ..., a_r \in A$, $b_1, ..., b_r \in B$ and $c_1, ..., c_r \in C$ decomposing $T(\alpha, \beta)$ as

$$T(\alpha, \beta) = \sum_{i=1}^{r} a_i(\alpha)b_i(\beta)c_i$$

for any $\alpha \in A^*$ and $\beta \in B^*$. The number $\mathrm{rk}(T)$ has also two additional interpretations.

- Considering T as an element of $A \otimes B \otimes C$ the rank r is the smallest number of rank one tensors in $A \otimes B \otimes C$ needed to span a linear space containing the point T. Let us rephrase this concept in the language of projective geometry. Consider the Segre embedding

$$\sigma_{1,1,1} : \mathbb{P}(A) \times \mathbb{P}(B) \times \mathbb{P}(C) \to \mathbb{P}(A \otimes B \otimes C),$$

  induced by $\mathcal{O}(1,1,1)$ and let $\Sigma_{1,1,1} = \sigma_{1,1,1}(\mathbb{P}(A) \times \mathbb{P}(B) \times \mathbb{P}(C))$ be the corresponding Segre variety. Then $\mathrm{rk}(T)$ is the smallest number of points $t_1, ..., t_r \in \Sigma_{1,1,1}$ in linear general position such that $[T] \in \langle t_1, ..., t_r \rangle$. In the language of secant varieties this means that $[T] \in \mathbb{S}ec_{r-1}(\Sigma_{1,1,1})^o$ but $[T] \notin \mathbb{S}ec_{r-2}(\Sigma_{1,1,1})^o$.
- Similarly, if we consider T as a linear map $A^* \to B \otimes C$ then $\mathrm{rk}(T)$ is the smallest number of rank one tensors in $B \otimes C$ need to span a linear space containing the linear space $T(A^*)$. As before we have a geometric counterpart. If we consider the Segre embedding

$$\sigma_{1,1} : \mathbb{P}(B) \times \mathbb{P}(C) \to \mathbb{P}(B \otimes C)$$

  and its image $\Sigma_{1,1}$ then $\mathrm{rk}(T)$ is the smallest number of points $t_1, ..., t_r \in \Sigma_{1,1}$ in linear general position such that $\mathbb{P}(T(A^*)) \subseteq \langle t_1, ..., t_r \rangle$.

The *border rank* of a bilinear map $T : A^* \times B^* \to C$ is the smallest natural number $r := \underline{\mathrm{rk}}(T)$ such that T is the limit of bilinear maps of rank r but is not a limit of tensors of rank s for any $s < r$. There is a geometric interpretations also for this notion: T has border rank r if $[T] \in \mathbb{S}ec_{r-1}(\Sigma_{1,1,1})$ but $[T] \notin \mathbb{S}ec_{r-2}(\Sigma_{1,1,1})$. Clearly $\mathrm{rk}(T) \geqslant \underline{\mathrm{rk}}(T)$.

Matrix Multiplication. Now, let us consider a special tensor. Given three vector spaces $L = \mathbb{C}^l, M = \mathbb{C}^m$ and $N = \mathbb{C}^n$ we have a matrix multiplication map

$$M_{m,n,l} : (N^* \otimes L) \times (L^* \otimes M) \to N^* \otimes M.$$

As a tensor $M_{m,n,l} = Id_N \otimes Id_M \otimes Id_L \in (N \otimes L^*) \otimes (L \otimes M^*) \otimes (N^* \otimes M)$, where $Id_N \in N^* \otimes N$ is the identity map. If $n = l$ the choice of a linear map $\alpha^0 : N \to L$ of maximal rank allows us to identify $N \cong L$. Then the multiplication map $M_{m,n,n} \in (N \otimes N^*) \otimes (N \otimes M^*) \otimes (N^* \otimes M)$ induces a linear map $N^* \otimes N \to (N^* \otimes M) \otimes (N^* \otimes M)^*$ which is an inclusion of Lie algebras

$$M : \mathfrak{gl}(N) \to \mathfrak{gl}(N^* \otimes M),$$

where $\mathfrak{gl}(N) \cong N^* \otimes N$ is the algebra of linear endomorphisms of N.

Matrix Equalities. The following lemmas are classical in linear algebra. However, for completeness we give a proof.

Lemma 1.1. *The determinant of a 2×2 block matrix is given by*

$$\det \begin{pmatrix} X & Y \\ Z & W \end{pmatrix} = \det(X) \det(W - ZX^{-1}Y),$$

where X is an invertible $n \times n$ matrix, Y is a $n \times m$ matrix, Z is a $m \times n$ matrix, and W is a $m \times m$ matrix.

Proof. The statement follows from the equality

$$\begin{pmatrix} X & Y \\ Z & W \end{pmatrix} \begin{pmatrix} -X^{-1}Y & Id_n \\ Id_m & 0 \end{pmatrix} = \begin{pmatrix} 0 & X \\ W - ZX^{-1}Y & Z \end{pmatrix}.$$

□

Lemma 1.2. *Let A be an $n \times n$ invertible matrix and U, V any $n \times m$ matrices. Then*

$$\det(A + UV^t) = \det(A)\det(Id + V^t A^{-1} U).$$

Proof. It follows from the equality

$$\begin{pmatrix} A & 0 \\ V^t & Id \end{pmatrix} \begin{pmatrix} Id & -A^{-1}U \\ 0 & Id + V^t A^{-1} U \end{pmatrix} \begin{pmatrix} Id & 0 \\ -V^t & Id \end{pmatrix} = \begin{pmatrix} A + UV^t & -U \\ 0 & Id \end{pmatrix}.$$

□

2. Landsberg - Ottaviani equations

In [LO] *J.M. Landsberg* and *G. Ottaviani* generalized Strassen's equations as introduced by *V. Strassen* in [S1]. We follow the exposition of [L, Section 2].
Let $T \in A \otimes B \otimes C$ be a tensor, and assume $b = c$. Let us consider T as a linear map $A^* \to B \otimes C$, and assume that there exists $\alpha \in A^*$ such that $T(\alpha) : B^* \to C$ is of maximal rank b. Via $T(\alpha)$ we can identify $B \cong C$, and consider $T(A^*) \subseteq B^* \otimes B$ as a subspace of the space of linear endomorphisms of B.
In [S1] *Strassen* considered the case $a = 3$. Let $\alpha^0, \alpha^1, \alpha^2$ be a basis of A^*. Assume that $T(\alpha^0)$ has maximal rank and that $T(\alpha^1), T(\alpha^2)$ are diagonalizable, commuting endomorphisms. Then $T(\alpha^1), T(\alpha^2)$ are simultaneously diagonalizable and both are a linear combination of the b rank one elements on the diagonal. Since $T(\alpha^0)$ has maximal rank $\underline{\mathrm{rk}}(T) \geqslant b$, and $\underline{\mathrm{rk}}(T) = b$ if and only if $[T(\alpha^1), T(\alpha^2)] = 0$. More generally $\underline{\mathrm{rk}}(T) \geqslant b + \mathrm{rank}[T(\alpha^1), T(\alpha^2)]/2$.
In particular the rank of the commutator $[M(\alpha^1), M(\alpha^2)]$ of $nm \times nm$ matrices is equal to m times the rank of the commutator $[\alpha^1, \alpha^2]$ of $n \times n$ matrices.
This equality reflects a general philosophy, that is to translate expressions in commutators of $\mathfrak{gl}_{n^2}$ into expressions in commutators in $\mathfrak{gl}_n$.
Now let us consider the case $a = 3, b = c$. Fix a basis a_0, a_1, a_2 of a A, and choose bases of B and C, so that elements of $B \otimes C$ can be written as matrices. Then we can write $T = a_0 \otimes X_0 - a_1 \otimes X_1 + a_2 \otimes X_2$, where the X_i are $b \times b$ matrices. Consider $T \otimes Id_A \in A \otimes B \otimes C \otimes A^* \otimes A = A^* \otimes B \otimes A \otimes A \otimes C$,

$$T \otimes Id_A = (a_0 \otimes X_0 - a_1 \otimes X_1 + a_2 \otimes X_2) \otimes (a^0 \otimes a_0 + a^1 \otimes a_1 + a^2 \otimes a_2)$$

and its skew-symmetrization in the A factor $T^1_A \in A^* \otimes B \otimes \bigwedge^2 A \otimes C$, which can be seen as a linear map

$$T^1_A : A \otimes B^* \to \bigwedge^2 A \otimes C$$

given by

$$a^1 X_0 (a_0 \wedge a_1) + a^2 X_0 (a_0 \wedge a_2) - a^0 X_1 (a_1 \wedge a_0) - a^2 X_1 (a_1 \wedge a_2) + a^0 X_2 (a_2 \wedge a_0) + a^1 X_2 (a_2 \wedge a_1).$$

In the basis a_0, a_1, a_2 of A and $a_0 \wedge a_1, a_0 \wedge a_2, a_1 \wedge a_2$ of $\bigwedge^2 A$ the matrix of T_A^1 is the following

$$Mat(T_A^1) = \begin{pmatrix} X_1 & -X_2 & 0 \\ X_0 & 0 & -X_2 \\ 0 & X_0 & -X_1 \end{pmatrix}$$

Assume X_0 is invertible and change bases such that it is the identity matrix. By Lemma 1.1, on the matrix obtained by reversing the order of the rows of $Mat(T_A^1)$, with

$$X = \begin{pmatrix} 0 & X_0 \\ X_0 & 0 \end{pmatrix},\ Y = \begin{pmatrix} -X_1 \\ -X_2 \end{pmatrix},\ Z = \begin{pmatrix} X_1 & -X_2 \end{pmatrix},\ W = 0$$

we get

$$\det(Mat(T_A^1)) = \det(X_1X_2 - X_2X_1) = \det([X_1, X_2]).$$

Now we want to generalize this construction as done in [LO]. We consider the case $a = 2p+1$, $T \otimes Id_{\bigwedge^p A} \in A \otimes B \otimes C \otimes \bigwedge^p A^* \otimes \bigwedge^p A = (\bigwedge^p A^* \otimes B) \otimes (\bigwedge^{p+1} A \otimes C)$, and its skew-symmetrization

$$T_A^p : \bigwedge^p A \otimes B^* \to \bigwedge^{p+1} A \otimes C.$$

Note that $\dim(\bigwedge^p A \otimes B^*) = \dim(\bigwedge^{p+1} A \otimes C) = \binom{2p+1}{p} b$. After choosing a basis $a_0, ..., a_{2p}$ of A we can write $T = \sum_{i=0}^{2p} (-1)^i a_i \otimes X_i$. The matrix of T_A^p with respect the basis $a_0 \wedge ... \wedge a_{p-1}, ..., a_{p+1} \wedge ... \wedge a_{2p}$ of $\bigwedge^p A$, and $a_0 \wedge ... \wedge a_p, ..., a_p \wedge ... \wedge a_{2p}$ of $\bigwedge^{p+1} A$ is of the form

$$Mat(T_A^p) = \begin{pmatrix} Q & 0 \\ R & \overline{Q} \end{pmatrix}$$

where the matrix is blocked $(\binom{2p}{p+1} b, \binom{2p}{p} b) \times (\binom{2p}{p+1} b, \binom{2p}{p} b)$, the lower left block is given by

$$R = \begin{pmatrix} X_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & X_0 \end{pmatrix}$$

and Q is a matrix having blocks $X_1, ..., X_{2p}$ and zero, while $\overline{Q}$ is the block transpose of Q except that if an index is even, the block is multiplied by -1.

Example 2.1. In the case $p = 2$ we have

$$\begin{aligned} T_A^2 = \ & (a^1 \wedge a^2)X_0(a_0 \wedge a_1 \wedge a_2) + (a^1 \wedge a^3)X_0(a_0 \wedge a_1 \wedge a_3) + (a^1 \wedge a^4)X_0(a_0 \wedge a_1 \wedge a_4) + \\ & (a^2 \wedge a^3)X_0(a_0 \wedge a_2 \wedge a_3) + (a^2 \wedge a^4)X_0(a_0 \wedge a_2 \wedge a_4) + (a^3 \wedge a^4)X_0(a_0 \wedge a_3 \wedge a_4) - \\ & (a^0 \wedge a^2)X_1(a_1 \wedge a_0 \wedge a_2) - (a^0 \wedge a^3)X_1(a_1 \wedge a_0 \wedge a_3) - (a^0 \wedge a^4)X_1(a_1 \wedge a_0 \wedge a_4) - \\ & (a^2 \wedge a^3)X_1(a_1 \wedge a_2 \wedge a_3) - (a^2 \wedge a^4)X_1(a_1 \wedge a_2 \wedge a_4) - (a^3 \wedge a^4)X_1(a_1 \wedge a_3 \wedge a_4) + \\ & (a^0 \wedge a^1)X_2(a_2 \wedge a_0 \wedge a_1) + (a^0 \wedge a^3)X_2(a_2 \wedge a_0 \wedge a_3) + (a^0 \wedge a^4)X_2(a_2 \wedge a_0 \wedge a_4) + \\ & (a^1 \wedge a^3)X_2(a_2 \wedge a_1 \wedge a_3) + (a^1 \wedge a^4)X_2(a_2 \wedge a_1 \wedge a_4) + (a^3 \wedge a^4)X_2(a_2 \wedge a_3 \wedge a_4) - \\ & (a^0 \wedge a^1)X_3(a_3 \wedge a_0 \wedge a_1) - (a^0 \wedge a^2)X_3(a_3 \wedge a_0 \wedge a_2) - (a^0 \wedge a^4)X_3(a_3 \wedge a_0 \wedge a_4) - \\ & (a^1 \wedge a^2)X_3(a_3 \wedge a_1 \wedge a_2) - (a^1 \wedge a^4)X_3(a_3 \wedge a_1 \wedge a_4) - (a^2 \wedge a^4)X_3(a_3 \wedge a_2 \wedge a_4) + \\ & (a^0 \wedge a^1)X_4(a_4 \wedge a_0 \wedge a_1) + (a^0 \wedge a^2)X_4(a_4 \wedge a_0 \wedge a_2) + (a^0 \wedge a^3)X_4(a_4 \wedge a_0 \wedge a_3) + \\ & (a^1 \wedge a^4)X_4(a_4 \wedge a_1 \wedge a_2) + (a^1 \wedge a^3)X_4(a_4 \wedge a_1 \wedge a_3) + (a^2 \wedge a^3)X_4(a_4 \wedge a_2 \wedge a_3) \end{aligned}$$

The matrix of T_A^2 is

$$Mat(T_A^2) = \begin{pmatrix} X_2 & -X_3 & X_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_1 & 0 & 0 & -X_3 & X_4 & 0 & 0 & 0 & 0 & 0 \\ 0 & X_1 & 0 & -X_2 & 0 & X_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & X_1 & 0 & -X_2 & X_3 & 0 & 0 & 0 & 0 \\ X_0 & 0 & 0 & 0 & 0 & 0 & -X_3 & X_4 & 0 & 0 \\ 0 & X_0 & 0 & 0 & 0 & 0 & -X_2 & 0 & X_4 & 0 \\ 0 & 0 & X_0 & 0 & 0 & 0 & 0 & -X_2 & X_3 & 0 \\ 0 & 0 & 0 & X_0 & 0 & 0 & -X_1 & 0 & 0 & X_4 \\ 0 & 0 & 0 & 0 & X_0 & 0 & 0 & -X_1 & 0 & X_3 \\ 0 & 0 & 0 & 0 & 0 & X_0 & 0 & 0 & -X_1 & X_2 \end{pmatrix}$$

If X_0 is the identity by Lemma 1.1 on $R = Id, Q$ and $\overline{Q}$ the determinant of $Mat(T_A^p)$ is equal to the determinant of

$$\begin{pmatrix} 0 & [X_1, X_2] & [X_1, X_3] & [X_1, X_4] \\ -[X_1, X_2] & 0 & [X_2, X_3] & [X_2, X_4] \\ -[X_1, X_3] & -[X_2, X_3] & 0 & [X_3, X_4] \\ -[X_1, X_4] & -[X_2, X_4] & -[X_3, X_4] & 0 \end{pmatrix}$$

In general the determinant of $Mat(T_A^p)$ is equal to the determinant of the $2pb \times 2pb$ matrix of commutators

$$\begin{pmatrix} 0 & X_{1,2} & X_{1,3} & X_{1,4} & \dots & X_{1,2p-1} & X_{1,2p} \\ -X_{1,2} & 0 & X_{2,3} & X_{2,4} & \dots & X_{2,2p-1} & X_{2,2p} \\ -X_{1,3} & -X_{2,3} & 0 & X_{3,4} & \dots & X_{3,2p-1} & X_{3,2p} \\ -X_{1,4} & -X_{2,4} & -X_{3,4} & 0 & \dots & X_{4,2p-1} & X_{4,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -X_{1,2p-1} & -X_{2,2p-1} & -X_{3,2p-1} & -X_{4,2p-1} & \dots & 0 & X_{2p-1,2p} \\ -X_{1,2p} & -X_{2,2p} & -X_{3,2p} & -X_{4,2p} & \dots & -X_{2p-1,2p} & 0 \end{pmatrix}$$

where $X_{i,j}$ denotes the commutator matrix $[X_i, X_j] = X_i X_j - X_j X_i$.

3. Key Lemma

We use the same notation of [L] throughout the text.

Lemma 3.1. [L1, Lemma 11.5.0.2] *Let V be a n-dimensional vector space and let $P \in S^d V^* \setminus \{0\}$ be a polynomial of degree $d \leqslant n-1$ on V. For any basis $\{v_1, ..., v_n\}$ of V there exists a subset $\{v_{i_1}, ..., v_{i_s}\}$ of cardinality $s \leqslant d$ such that $P_{|\langle v_{i_1}, ..., v_{i_s} \rangle}$ is not identically zero.*

Proof. Let $x = \sum_{i=1}^n x_i v_i$ be an element of U and consider $P(x)$ as a polynomial in $x_1, ..., x_n$. For instance take the first non-zero monomial appearing in $P(x)$. Since it can involve at most d of the x_i's the polynomial P restricted to the span of the corresponding v_i's is not identically zero. □

Lemma 3 says, for instance, that a quadric surface in $\mathbb{P}^3$ can not contain six lines whose pairwise intersections span $\mathbb{P}^3$. Note that as stated Lemma 3 is sharp. For example the polynomial $P(x, y, z, w) = xy$ vanishes on the points $[1 : 0 : 0 : 0], ..., [0 : 0 : 0 : 1] \in \mathbb{P}^3$.

Lemma 3.2. *Let $A = N^* \otimes L$, where $l = n$. Given any basis of A, there exists a subset of at least $n^2 - (2p+3)n$ basis vectors, and elements $\alpha^0, \alpha^1, \ldots, \alpha^{2p}$ of A^*, such that*

- *α^0 is of maximal rank, and thus may be used to identify $L \simeq N$ and A as a space of endomorphisms. (I.e. in bases α^0 is the identity matrix.)*
- *Choosing a basis of L, so the α^j become $n \times n$ matrices, the size $2pn$ block matrix whose (i,j)-th block is $[\alpha^i, \alpha^j]$ has non-zero determinant, and*
- *The subset of $n^2 - (2p+3)n$ basis vectors annihilate $\alpha^0, \alpha^1, \ldots, \alpha^{2p}$.*

Proof. Let $\mathcal{B}$ be a basis of A, and consider the polynomial $P_0 = \det_n$. By Lemma 3.1 we get a subset S_0 of at most n elements of $\mathcal{B}$ and $\alpha^0 \in S_0$ with $\det_n(\alpha^0) \neq 0$. Now, via the isomorphism $\alpha^0 : L \to N$ we are allowed to identify $A = \mathfrak{gl}(L)$ as an algebra with identity element α^0. So, from now on, we work with $\mathfrak{sl}(L) = \mathfrak{gl}(L)/\langle \alpha^0 \rangle$ instead of $\mathfrak{gl}(L)$.
Following the proof of [L, Lemma 4.3] we consider $v_{1,0}, ..., v_{2p,0} \in \mathfrak{sl}(L)$ be linearly independent and not equal to any of the given basis vectors, and work locally on an affine open neighbourhood $\mathbb{V} \subset G(2p, \mathfrak{sl}(L))$ of $E_0 = \langle v_{1,0}, ..., v_{2p,0} \rangle$. We extend $v_{1,0}, ..., v_{2p,0}$ to a basis $v_{1,0}, ..., v_{2p,0}, w_1, ..., w_{n^2-2p-1}$ of $\mathfrak{sl}(L)$, and take local coordinates (f_s^μ) with $1 \leqslant s \leqslant 2p$, $1 \leqslant \mu \leqslant n^2 - 2p - 1$, on V, so that $v_s = v_{s,0} + \sum_{\mu=1}^{n^2-2p-1} f_s^\mu w_\mu$.
We denote $v_{i,j} = [v_i, v_j]$ and let us define

$$A_{i,i+1} = \begin{pmatrix} 0 & v_{i,i+1} \\ -v_{i,i+1} & 0 \end{pmatrix}$$

for $i = 1, \ldots, 2p$ and let A be the following diagonal block matrix

$$A = \operatorname{diag}(A_{1,2}, A_{3,4}, \ldots, A_{2p-3,2p-2}, Id_{2n \times 2n})$$

which is a squared matrix of order $4pn$. Consider the $4pn \times 4pn$ matrix

$$M = \begin{pmatrix} 0 & v_{1,2} & v_{1,3} & v_{1,4} & \ldots & v_{1,2p-1} & v_{1,2p} \\ -v_{1,2} & 0 & v_{2,3} & v_{2,4} & \ldots & v_{2,2p-1} & v_{2,2p} \\ -v_{1,3} & -v_{2,3} & 0 & v_{3,4} & \ldots & v_{3,2p-1} & v_{3,2p} \\ -v_{1,4} & -v_{2,4} & -v_{3,4} & 0 & \ldots & v_{4,2p-1} & v_{4,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} & -v_{3,2p-1} & -v_{4,2p-1} & \ldots & 0 & v_{2p-1,2p} \\ -v_{1,2p} & -v_{2,2p} & -v_{3,2p} & -v_{4,2p} & \ldots & -v_{2p-1,2p} & 0 \end{pmatrix}$$

The polynomial $\det_{4pn \times 4pn}(M)$ is not identically zero on $G(2p, \mathfrak{sl}(L))$, so it is not identically zero on $\mathbb{V}$. Furthermore we can write $M = A + UId_{4pn \times 4pn}$, where

$$U = \begin{pmatrix} 0 & 0 & v_{1,3} & v_{1,4} & \ldots & v_{1,2p-1} & v_{1,2p} \\ 0 & 0 & v_{2,3} & v_{2,4} & \ldots & v_{2,2p-1} & v_{2,2p} \\ -v_{1,3} & -v_{2,3} & 0 & 0 & \ldots & v_{3,2p-1} & v_{3,2p} \\ -v_{1,4} & -v_{2,4} & 0 & 0 & \ldots & v_{4,2p-1} & v_{4,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} & -v_{3,2p-1} & -v_{4,2p-1} & \ldots & -Id_{n \times n} & v_{2p-1,2p} \\ -v_{1,2p} & -v_{2,2p} & -v_{3,2p} & -v_{4,2p} & \ldots & -v_{2p-1,2p} & -Id_{n \times n} \end{pmatrix}$$

By Lemma 1.2 we have

$$\det(M) = \det(A) \det(Id + A^{-1}U) = \det([v_1, v_2])^2 \ldots \det([v_{2p-3}, v_{2p-2}])^2 \det(Id + A^{-1}U).$$

The entries of the $n \times n$ matrices $[v_k, v_{k+1}]$ are quadratic in the f_s^μ's, so the polynomials $\det([v_k, v_{k+1}])$ have degree $2n$, and

$$P_1 = \det([v_1, v_2])^2 \dots \det([v_{2p-3}, v_{2p-2}])^2 = (\det([v_1, v_2]) \dots \det([v_{2p-3}, v_{2p-2}]))^2$$

is a polynomial of degree $4n(p-1)$. Since P_1 is a square, we can consider the polynomial $\widetilde{P}_1 = \det([v_1, v_2]) \dots \det([v_{2p-3}, v_{2p-2}])$ which has degree $2n(p-1)$. Applying Lemma 3.1 to $\widetilde{P}_1$ we find a subset S_1 of at most $2n(p-1)$ elements of our basis such that $\widetilde{P}_1$, and hence P_1, is not identically zero on $\langle S_1 \rangle$.
Now, let us fix some particular value of the coordinates f_s^μ such that on the corresponding matrices $\overline{v}_1, ..., \overline{v}_{2p-2}$ the matrix A is invertible. For these values the expression $\det(Id + A^{-1}U)$ make sense. Let us consider the matrix

$$Id + A^{-1}U = \begin{pmatrix} Id & 0 & -v_{1,2}^{-1}v_{2,3} & -v_{1,2}^{-1}v_{2,4} & \dots & -v_{1,2}^{-1}v_{2,2p-1} & -v_{1,2}^{-1}v_{2,2p} \\ 0 & Id & v_{1,2}^{-1}v_{1,3} & v_{1,2}^{-1}v_{1,4} & \dots & v_{1,2}^{-1}v_{1,2p-1} & v_{1,2}^{-1}v_{1,2p} \\ v_{3,4}^{-1}v_{1,4} & v_{3,4}^{-1}v_{2,4} & Id & 0 & \dots & -v_{3,4}^{-1}v_{4,2p-1} & -v_{3,4}^{-1}v_{4,2p} \\ -v_{3,4}^{-1}v_{1,3} & -v_{3,4}^{-1}v_{2,3} & 0 & Id & \dots & v_{3,4}^{-1}v_{3,2p-1} & v_{3,4}^{-1}v_{3,2p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} & -v_{3,2p-1} & -v_{4,2p-1} & \dots & 0 & v_{2p-1,2p} \\ -v_{1,2p} & -v_{2,2p} & -v_{3,2p} & -v_{4,2p} & \dots & -v_{2p-1,2p} & 0 \end{pmatrix}$$

By Lemma 1.1 on $Id + A^{-1}U$ with

$$X = \begin{pmatrix} Id & 0 \\ 0 & Id \end{pmatrix}, \; Y = \begin{pmatrix} -v_{1,2}^{-1}v_{2,3} & -v_{1,2}^{-1}v_{2,4} & \dots & -v_{1,2}^{-1}v_{2,2p-1} & -v_{1,2}^{-1}v_{2,2p} \\ v_{1,2}^{-1}v_{1,3} & v_{1,2}^{-1}v_{1,4} & \dots & v_{1,2}^{-1}v_{1,2p-1} & v_{1,2}^{-1}v_{1,2p} \end{pmatrix},$$

$$Z = \begin{pmatrix} v_{3,4}^{-1}v_{1,4} & v_{3,4}^{-1}v_{2,4} \\ -v_{3,4}^{-1}v_{1,3} & -v_{3,4}^{-1}v_{2,3} \\ \vdots & \vdots \\ -v_{1,2p-1} & -v_{2,2p-1} \\ -v_{1,2p} & -v_{2,2p} \end{pmatrix}, \; W = \begin{pmatrix} Id & 0 & \dots & -v_{3,4}^{-1}v_{4,2p-1} & -v_{3,4}^{-1}v_{4,2p} \\ 0 & Id & \dots & v_{3,4}^{-1}v_{3,2p-1} & v_{3,4}^{-1}v_{3,2p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -v_{3,2p-1} & -v_{4,2p-1} & \dots & 0 & v_{2p-1,2p} \\ -v_{3,2p} & -v_{4,2p} & \dots & -v_{2p-1,2p} & 0 \end{pmatrix}$$

we get $\det(Id + A^{-1}U) = \det(W - ZY)$. Note that the coordinates f_s^μ appear in the terms indexed by $2p-1$ and $2p$, while all the other terms are constant once we fixed $\overline{v}_1, ..., \overline{v}_{2p-2}$. Then $P_2 = \det(W - ZY)$ is a polynomial of degree $4n$. By Lemma 3.1 we find a subset S_2 of at most $4n$ elements of the basis $\mathcal{B}$ such that P_2 is not identically zero on $\langle S_2 \rangle$.
Summing up we found a subset S of at most $n + 2n(p-1) + 4n = (2p+3)n$ elements of $\mathcal{B}$ such that $\det(M)$ is not identically zero on $\langle S \rangle$. □

Remark 3.3. In [L, Lemma 4.3] the author proved the analogous statement for $n^2 - (4p+1)n$.

We are ready to prove Theorem 0.1 following the proof of [L, Theorem 1.2].

Proof. Let φ be a decomposition of the matrix multiplication tensor $M_{n,n,m}$ as sum of $r = \mathrm{rk}(M_{n,n,m})$ rank one tensors. Recall that the left kernel of a bilinear map $f : V \times U \to W$ is defined as $\mathrm{Lker}(f) = \{v \in V \mid f(v,u) = 0 \; \forall \; u \in U\}$. Since $\mathrm{Lker}(M_{n,n,m}) = 0$, that is for any $\alpha \in A^* \setminus \{0\}$, there exists $\beta \in B^*$ such that $M_{n,n,m}(\alpha, \beta) \neq 0$ we can write $\varphi = \varphi_1 + \varphi_2$ with $\mathrm{rk}(\varphi_1) = n^2$, $\mathrm{rk}(\varphi_2) = r - n^2$ and $\mathrm{Lker}(\varphi_1) = 0$.

The n^2 elements of A^* appearing in φ_1 form a basis of A^*. By Lemma 3.2 there exists a subset of $n^2-(2p+3)n$ of them annihilating a maximal rank element α^0 and some $\alpha^1, ..., \alpha^{2p}$ such that, choosing bases, the determinant of the matrix $([\alpha^i, \alpha^j])$ is non-zero.
Let ψ_1 be the sum of all monomial in φ_1 whose terms in A^* annihilate $\alpha^0, ..., \alpha^{2p}$. By Lemma 3.2 there exists at least $n^2 - (2p+3)n$ of them. Then $\mathrm{rk}(\psi_1) \geqslant n^2 - (2p+3)n$. Furthermore consider $\psi_2 = \varphi_1 - \psi_1 + \varphi_2$ so that $\varphi = \psi_1 + \psi_2$ and the terms appearing in ψ_2 does not annihilate $\alpha^0, ..., \alpha^{2p}$.
Let $A^{'} = \langle \alpha^0, ..., \alpha^{2p} \rangle \subseteq A^*$. Again by Lemma 3.2 the determinant of the linear map $M_{n,n,m|A'\otimes B^*\otimes C^*} : \bigwedge^p A^{'} \otimes B^* \to \bigwedge^{p+1} A^{'} \otimes C$ is non-zero. Then $\underline{\mathrm{rk}}(\varphi_2) \geqslant nm\frac{2p+1}{p+1} = \dim(\bigwedge^p A^{'} \otimes B^*)$. We conclude that

$$\mathrm{rk}(\varphi) = \mathrm{rk}(\varphi_1) + \mathrm{rk}(\varphi_2) \geqslant n^2 - (2p+3)n + nm\frac{2p+1}{p+1} = (1 + \frac{p}{p+1})nm + n^2 - (2p+3)n.$$

This concludes the proof of (0.1).
To prove the other assertions, let us consider the function $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ defined by $f(p) = (3 - \frac{1}{p+1})n^2 - (2p+3)n$. The first derivative is $f'(p) = \frac{1}{(p+1)^2}n^2 - 2n$, which vanishes in $p = \sqrt{\frac{n}{2}} - 1$. Moreover $f''(p) = -\frac{2}{(p+1)^3}n^2 < 0$, hence $p = \sqrt{\frac{n}{2}} - 1$ is the maximum of f.
Then the bound (0.2) is maximized for $p = \lceil \sqrt{\frac{n}{2}} - 1 \rceil$ or $p = \lfloor \sqrt{\frac{n}{2}} - 1 \rfloor$, depending on the value of n.
If $(\sqrt{\frac{n}{2}} - 1) - \lfloor \sqrt{\frac{n}{2}} - 1 \rfloor \geqslant \frac{1}{2}$ we may consider $p = \lceil \sqrt{\frac{n}{2}} - 1 \rceil$. In this case $\sqrt{\frac{n}{2}} - 1 \leqslant p \leqslant \sqrt{\frac{n}{2}} - \frac{1}{2}$, and we get $f\big(\lceil \sqrt{\frac{n}{2}} - 1 \rceil\big) \geqslant \lceil f \rceil(n) := 3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 2n$.
If $(\sqrt{\frac{n}{2}} - 1) - \lfloor \sqrt{\frac{n}{2}} - 1 \rfloor < \frac{1}{2}$ we consider $p = \lfloor \sqrt{\frac{n}{2}} - 1 \rfloor$. Then $\sqrt{\frac{n}{2}} - \frac{3}{2} \leqslant p \leqslant \sqrt{\frac{n}{2}} - 1$, and we have $f\big(\lfloor \sqrt{\frac{n}{2}} - 1 \rfloor\big) \geqslant \lfloor f \rfloor(n) := (3 - \frac{2\sqrt{2}}{2n - \sqrt{2}})n^2 - \sqrt{2}n^{\frac{3}{2}} - n$.
Finally to prove (0.3) it is enough to observe that both $\lceil f \rceil(n)$ and $\lfloor f \rfloor(n)$ are greater than $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$. □

Acknowledgements. Both the authors were introduced to this topic by *J.M. Landsberg* during the Summer School *"Tensors: Waring problems and Geometric Complexity Theory"* held in Cortona in July 2012. We would like to thank *J.M. Landsberg* and *M. Mella* for their beautiful lectures and all the participants for the stimulating atmosphere. We thank primarily *J.M. Landsberg* for suggesting us the problem, for his interest and his generous suggestions.

References

[B] M. Bläser, *A $\frac{5}{2}n^2$ lower bound for the rank of $n \times n$-matrix multiplication over arbitrary fields*, 440th Annual Symposium on Foundations of Computer Science (New York, 1999), IEEE Computer Soc, Los Alamitos, CA, 1999, pp. 45-50, MR MR1916183.

[L] J.M. Landsberg, *New lower bounds for the rank of matrix multiplication*, arXiv:1206.1530.

[L1] J.M. Landsberg, *Tensors: geometry and applications*, Graduate Studies in Mathematics, vol. 128, American Mathematical Society, Providence, RI, 2012, MR 2865915.

[LO] J.M. Landsberg, G. Ottaviani, *New lower bounds for the border rank of matrix multiplication*, arXiv:1112.6007.

[S] V. Strassen, *Gaussian Elimination is not Optimal*, Numer. Math. 13, p. 354-356, 1969.

[S1] V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl. 52/53(1983), 645-685. MR 85b:15039.

Alex Massarenti, SISSA, via Bonomea 265, 34136 Trieste, Italy
E-mail address: alex.massarenti@sissa.it

Emanuele Raviolo, Università di Pavia, via Ferrata 1, 27100 Pavia, Italy
E-mail address: emanuele.raviolo@unipv.it